The evolution of metabolic strategies

Wortel, M.T.

2015

8 General Discussion

The evolution of metabolic strategies

In this thesis we characterized the properties of the expected metabolic strategies for constant environments, in which the selection is on the specific growth rate. One way to achieve a high specific growth rate is the use of efficient enzymes—those that maximize the flux per unit enzyme. In chapter 3 we show that efficient transporters do not necessarily have a high affinity, which first intuition would suggest. Actually, at high substrate concentrations, low affinity transporters achieve a higher uptake flux and also increase the flux through other parts of metabolism, such as glycolysis in *S. cerevisiae*.

When we know the set of available enzymes and their properties, we can show which enzymes need to be expressed to achieve maximal specific growth rate (chapter 4), namely enzymes that catalyse reactions that form an elementary flux mode (EFM). EFMs were previously defined to characterize solutions spaces of stoichiometric networks (Schuster and Hilgetag, 1994) and interestingly, now turn out to have a rather different application for kinetic optimization.

The availability of kinetic information of enzymes is still limited, and therefore flux balance analysis (FBA) (Orth *et al*, 2010), which optimizes growth combined with constraints derived from experimental data, is more practical. FBA with only a single flux constraint optimizes the yield, relevant in spatial environments and environments with privatised resources. FBA leads to large sets of solutions that all achieve optimal constrained growth rate, which are hard to visualize and understand. In chapter 5 we developed a tool to enumerate all pathways in this optimal space, showing the metabolic capacities in the optimal space in the same way that EFMs show the metabolic capacities in the total solution space. Because of the large solution space, we expect a big influence of secondary objectives on the metabolic behaviour. We showed that two secondary objectives, minimization of active reactions and enzyme costs, reduce the solution space to respectively a smaller space and a single solution.

To obtain a more complete understanding of the whole system, we modeled the interaction of the metabolic strategy with the environment using a combined metabolic and population model. We choose the chemostat environment, which is constant when we look at the population dynamics, but does change slowly at evolutionary timescales. In contrast with what was previously reported, namely selection for increased substrate affinity (a.o. Brown *et al* (1998)), the selection pressure in the chemostat is on growth rate (chapter 6). An inhibiting product of the optimal specific rate EFM, such as in ethanol production in *S. cerevisiae*, could lead to interesting evolutionary dynamics. When the concentration of inhibiting product is high, a strategy not producing this product—not expressing the enzyme targeted by the inhibitor—is selected for. Conversely, when the concentration of inhibiting product is low, a strategy producing this product is selected for. This density dependent selection can lead to the evolution and coexistence of two different metabolic strategies in the chemostat.

Dynamic environments can select for metabolic strategies that would cause a disadvantage under constant conditions (chapter 7). When a side branch of glycolysis that goes towards trehalose—a metabolite protecting the cell from stresses such as freezing and dehydration (François and Parrou, 2001)—does not function properly, a destructive phenotype emerges

after a glucose pulse. Transient futile cycling that wastes energy, something we would never expect in constant environments, turns out to be a sophisticated mechanism to guarantee a healthy metabolic state after sudden changes in nutrient levels. It may be a additional function for a stress protection mechanism, but perhaps steering the glycolysis was the original function of this branch. Trehalose itself protects against stress but does not have a direct function in the glucose pulse response, however, there might have been a cycle without trehalose first.

The influence of enzyme kinetics

The importance of enzyme kinetics for the choice of metabolic strategies has a fundamental as well as practical interest. Kinetic data are not yet and might never be available at large scale, highlighting the importance of knowing for which questions we do, and for which we do not have to include the kinetics. The optimization of yield (chapter 5) and the possible metabolic structures that optimize the specific growth rate (chapter 4) do not require any kinetic information. However, the specific growth rate, the selection pressure in constant environments, depends on the kinetics. The set and level of enzymes that optimize growth rate can be different from the one that optimizes yield, and cannot be predicted from a stoichiometric network only, therefore we need kinetics to unravel the evolution in constant environments.

Kinetic properties of transporter enzymes (chapter 3) might be related to metabolic switches (chapter 2). Low affinity transporters can carry a high flux at high extracellular substrate concentrations and are less sensitive to a high intracellular nutrient concentration, because of their low affinity. Low affinity enzymes of consecutive steps can carry a high flux when the intracellular concentration is high. High extracellular concentrations can therefore select for a metabolic strategy that is otherwise beneficial but includes low affinity enzymes. These changes in enzymes with the substrate concentration can lead to a metabolic switch, where at low extracellular concentrations the mode with high affinity enzymes is used, and at high extracellular concentrations the mode with low affinity enzymes (chapter 4). In accordance, Crabtree negative yeasts, which have only high affinity transporters (Van Urk *et al*, 1989), do not have a metabolic switch; the lack of low affinity transporters might prevent the low affinity pathway from becoming optimal, even when extracellular nutrient concentrations are high.

Youk and Van Oudenaarden (2009) and Huberts *et al* (2012) suggest that the metabolic switch in *S. cerevisiae* is mainly mediated by the uptake rate and Huberts *et al* (2012) argues that this rate is signaled by measuring the intracellular fluxes. Huberts *et al* (2012) suggests that cells infer intracellular fluxes from some concentrations of internal metabolites, since these are often correlated. In the previous paragraph, we suggest that selection for the metabolic switch is related to intracellular metabolite concentrations, providing an interesting link between both theories.

Product inhibition results from the binding of a product to the enzyme by which it is produced, preventing the enzyme from performing its function. This inhibition affects metabolic network behaviour in interesting ways, and therefore the evolution of metabolic strategies. In chapter 7 we show that the addition of ethanol to the medium improves the survival of the tpsΔ mutants after a glucose pulse. We propose an explanation, motivated by the results of our kinetic model,

that the product inhibition of ethanol on alcohol dehydrogenase changes the redox state of the cells. The increased NADH stimulates the glycerol branch which increases the chance to make it through the transition. In chapter 6 we show that this same inhibition of ethanol can lead to the divergent evolution of metabolic strategies leading to subpopulations in the chemostat. It is one of many illustrations how a certain strategy or property—in this case ethanol production and inhibition—affects different aspects of fitness.

In constant conditions we do not expect futile cycling and other metabolic structures that do not use the enzymes efficiently. However, we have shown in chapter 7 that futile cycling is necessary to steer the metabolism towards a stable state after a glucose pulse. The kinetics of this cycle—namely inhibition by phosphate—are such that the flux through the cycle is very low once a stable state has been reached and the environment remains constant. In this manner the kinetic properties of the enzymes are smartly used to reduce the negative effects of this safety mechanism under constant conditions, and come close to the metabolic strategy that optimises growth rate.

Observed metabolic strategies

Our theories for optimal metabolic strategies can lead to hypotheses about metabolic strategies in different environments. In the end we would like to use them to predict the evolution of metabolism under controlled conditions. The other way around, we can use the theory to deduce the previous selection pressures and environments from the observed metabolic strategies. However, real life is obviously even more complex than our models. The main challenges for the link between model and experiment include: (1) the lack of data, such as kinetic data for growth rate optimization. (2) The scarcity of constant environments; most environments exert a combination of selection pressures and we as yet have no complete theory that links metabolic strategies to combinations of selection pressures (for a study with different selection pressures combined with FBA see Schuetz *et al* (2012)). (3) The duration of adaptation; whether cells have been evolved in the same environment to adapt close to optimal strategies and how much deviation from optimal strategies we can expect due to e.g. genetic drift. Can we still deduce the selection pressure if cells deviate slightly from optimality, perhaps in different ways within a population? (4) Characterizations at the population level will not be accurate when there are phenotypic or genetic subpopulations, which have been found in well mixed environments (chapter 7 and Solopova *et al* (2014); Beaumont *et al* (2009)).

Experimental manipulations of enzyme levels

That enzymes are costly can be tested by introducing useless enzyme—dummy protein—and observing the effect on the growth rate. Such experiments in *E. coli* have shown that dummy protein production reduces the growth rate, confirming that enzyme production is costly (Dekel and Alon (2005) and references in Shachrai *et al* (2010)). Lang *et al* (2009) reduced the protein burden in *S. cerevisiae* by making strains that do not produce enzymes for mating. They found that strains with lower protein burden had a higher growth rate when reproducing asexually.

When enzyme expression is costly, enzyme levels need to be tuned precisely to obtain the maximal growth rate (Berkhout *et al* (2013a) and chapter 4). The most straightforward method

to test this hypothesis is to experimentally adjust the enzyme levels and observe whether this decreases the growth rate. Bollenbach *et al* (2009) show that wild type *E. coli* cells have optimal expression of ribosomal proteins. Solem *et al* (2008) introduced a β-galoctidase from *Streptococcus thermophilus* in *L. lactis* at different levels and observed an intermediate optimal level of β-galoctidase. They could not measure the wildtype expression level, but did observe that the growth rate at the optimal level of β-galoctidase was similar to the wild-type growth rate.

Producing only the enzymes that are strictly necessary reduces the flexibility of a cell, for example if only transporters are expressed for one of the nutrients that is available, cells cannot grow when that nutrient runs out. Especially when enzymes have to be present only in low amounts to sustain growth under conditions that would, without the enzyme, not allow growth, we might expect enzymes to deviate from their optimal abundance. On the other hand, ribosomes are very large enzymes that are always needed. As expected, ribosome expression is tightly linked to growth rate, increasing linearly with the growth rate (Scott *et al*, 2010). The linear increase results from a growth rate optimization (Bosdriesz *et al*, 2015; Scott *et al*, 2014).

A sudden change, such as a sudden high production of dummy protein, leaves cells badly adapted to the situation. In time, cells can adjust to high dummy protein levels; Shachrai *et al* (2010) showed a reduction in cost after several generations of unneeded protein producing *E. coli* in exponential phase.

Enzymes are costly because they make the cell bigger; more biomass has to be produced to produce a new cell. However, Stoebel *et al* (2008) found that the fitness of an *E. coli* strain with dummy protein—the lac operon under glucose conditions—did not change when the proteins were recycled (and thus the protein burden was less but the production costs were in effect). For these proteins most of the costs come from the expression process rather than the used amino acids. It would be very interesting to investigate how these costs differ among enzymes, e.g. the relative costs for metabolic proteins and ribosomes.

Evolution experiments at constant conditions

In the laboratory, the chemostat environment approaches constant conditions, although not at evolutionary timescales. The chemostat has been used for several evolution experiments with *S. cerevisiae* and *E. coli* (Helling *et al*, 1987) (summarized in Gresham and Hong (2014)). All experiments with glucose limitations report a decrease in residual glucose (Ferea, 1999; Jansen *et al*, 2005; Gresham *et al*, 2008; Mashego *et al*, 2005) and several report a shift to higher affinity transport.

Any adaptation to the conditions in a chemostat will eventually lead to less favorable external conditions—decrease of the residual substrate or increase of inhibiting compounds such as ethanol. Selection on growth rate at the prevailing conditions transiently increases the biomass concentration and, in turn, the substrate concentration decreases to a new steady state at a lower residual substrate and/or higher inhibitor concentration. The increase of high affinity transport and shift to respirative metabolism can be the result of two phenomena: (1) improvements at the original substrate concentration and, as a consequence, a decrease in residual glucose or (2) first a decrease in residual glucose concentration by a different cause and then adaptations to the decreased concentrations. The question to distinguish between those two ef-

fects has not yet been addressed, but likely both effects will play a role. An experimental set-up that will keep the nutrient concentration constant at evolutionary timescales—a "glucostat"—could be used to distinguish between these two effects.

We predict a single EFM to be optimal for specific growth rate as opposed to a mixed strategy, at every substrate concentration (chapter 4). This means that for micro-organisms that shift their metabolism with the growth rate (chapter 2), we would expect a more or less abrupt switch from one mode of metabolism to the other as the substrate increases. Microorganisms can be observed in chemostats at different dilution rates; *S. cerevisiae* populations show a mixture of respiration and fermentation at a range of growth rates in the chemostat. Although this does not concur with our theory for an abrupt switch, this suboptimal adaptation can be explained by the natural environment of yeast; yeast cells are—in their evolutionary history—probably more adapted to excess glucose concentrations (bursting fruits) and low glucose concentrations (transitions) than to relatively high but limiting glucose concentrations at high growth rates in the chemostat.

Experimental results that can give us some insight in the advantage of a discrete switch include adaptation experiments at low growth rates, showing a shift from mixed to fully respiratory metabolism and an increased yield (Jansen *et al*, 2005; Brown *et al*, 1998; Ferea, 1999). Unfortunately, these experiments are not performed at higher growth rates. Although no conclusive experiments have been performed to discover if a discrete or a gradual switch is evolutionary favorable, Metris *et al* (2014) observed an abrupt switch from aerobic to fermentative pathways in *E. coli* at increasing concentrations of NaCl and Ibarra *et al* (2002) observed a shift towards more overflow metabolism at constant high glucose concentrations (in batch conditions).

Evolution experiments under dynamic conditions

Long term evolution with micro-organisms in batch is convenient (Elena and Lenski, 2003) and is therefore the most common experimental set-up (e.g. Lenski and Travisano (1994); Bongers *et al* (2003); Bachmann *et al* (2012)). However, dynamics in this environment, the alternation between high and low nutrients, lead to evolution of increased maximum growth rate and reduced lag phase (Vasi *et al*, 1994). The scarcity of nutrients at some stages can lead to the unexpected evolution of traits such as the growth on a new compound, as in citrate growth in the long term *E. coli* cultures (Blount *et al*, 2008).

The periodic environments with changing nutrient availability, biomass concentrations, and waste product concentrations in batch cultures, make it difficult to elucidate the dominant selection pressure. The selection pressure could be on high maximal growth rate, survival in stationary phase, and dealing with sudden nutrient pulses. The last selection pressure could lead to more flexibility and for example futile cycles (chapter 7).

Outlook

Developments in high-throughput techniques and computational methods, to which the work in this thesis contributed, open up many interesting directions for further research. Future research will expand the specific growth rate optimization (presented in chapter 4) to genome scale, which requires an improvement of methods, and a breakthrough in the high throughput generation of

kinetic data. The last might be a bottleneck, but perhaps new techniques combined with smart ideas of generating reasonable kinetic information from attainable data will bring a surprise. An approximation with non-saturating kinetics, such as linlog (Visser and Heijnen, 2003) and generalised mass action (Savageau, 1969), will not be useful for optimization purposes. These methods are reasonably accurate at a range of metabolite values, but can be very different for extreme values, and with optimizations the complete range of metabolite values is of importance.

Subpopulations are observed in experimental populations of microorganisms (Solopova *et al* (2014); Beaumont *et al* (2009) and chapter 7) and are predicted by theoretical studies (chapter 6). Methods to detect subpopulations require analysis on single cell level, such as fluorescence-activated cell sorting and microscopy methods. Methods determining the strategies of single cells can distinguish if observed mixed strategies are population or single cell properties, and if the subpopulations predicted in chapter 6 indeed exist.

Biological variability can originate from genetic or phenotypic variability. The genetic variability drives the evolution, but phenotypic variability can also have profound effects (chapter 7). Genetic variability can arise due to mutations and horizontal gene transfer (Thomas and O'Shea, 2005), but also from whole genome duplications (Oud *et al*, 2013; Conant and Wolfe, 2007). An important consideration is what possible changes we can expect in the metabolic network, a slight shift in strategy due to a mutation in the affinity of a regulator or a complete change due to a mutation that stops the expression of a gene completely. In chapter 6 we allowed for both types of mutations, with a higher chance for smaller changes. In the future, information about the origin of the variation in genotypes and the intermediate phenotypes can be used to create fitness landscapes (De Visser and Krug, 2014) and predict evolutionary paths in controlled settings.

In this thesis we fully focussed on micro-organisms but some of our results will also be applicable to different types of cells. Multicellular organisms offer the challenge that different cells will have different objectives, and our results, such as the selection on specific growth rate in constant conditions, do not hold in general. Selection will be on the reproduction of the organism as a whole, and experimental evolution will be more difficult due to long generation times. Only for cells of multicellular organisms that behave and are selected in a similar way as microorganisms, such as cancer cells and cell lines, can our theory be applied. Intriguingly, such cells also exhibit a metabolic shift (chapter 2). To predict metabolic strategies of multicellular organisms is still a future challenge.

Although we made many interesting discoveries, it is still a long way to the prediction of evolution, even under tightly controlled conditions. We mentioned the expansion of the theory to genome-scale, more large-scale data generation and the inclusion of genetic trajectories. Also, to bring the theory further, we propose a shift from evolution experiments of an explorative nature to experiments that are more focussed on outcomes that can be compared to the theory, such as experiments under constant nutrients at evolutionary scale and parallel experiments under different conditions.